BIG DATA

LOUISE MAYNARD-ATEM

What's the Big (data) Deal?

In my introductory article I mentioned that I would be writing bite-sized tutorials on various essential techniques within the O.R. toolkit that would prove particularly useful to those of us at the start or in the early stages of our careers. Initially I thought I might kick-off with multi-criteria decision analysis (MCDA), explaining the basics of the technique and giving examples of how it is used by operational researchers in their working lives; however, since attending the recent annual Government OR conference, I have decided instead to talk about big data/predictive analytics – what are they, what do they mean to the O.R. community and how should we, as early operational researchers, try and incorporate them into our toolkit of skills?

So what actually is big data...?

It is perhaps a little self-explanatory but big data is just that – an informal definition is datasets so large in size that they cannot be easily captured or processed using regular software in a reasonable amount of time. The formal definition, first proposed by Doug Laney in 2001, described exponential data growth as being three-dimensional i.e. increasing volume (amounts of data), velocity (speed of data flow) and variety (taking into account the range of types and sources of data). In 2012, the definition was updated (by Gartner Inc.) as follows:

'Big data is high volume, high velocity, and/or high variety information assets that require new forms of processing to enable enhanced decision making, insight discovery and process optimisation.'

This definition is now widely used and commonly referred to as the 3V's, with some organisations adding a fourth V – representing veracity (how trustworthy is the data).

Predictive analytics follows on almost automatically from any conversation on big data and can be formally defined as an area of data mining that deals with extracting information from data and using it to identify trends in activity and behaviour. It is hoped that by applying predictive analytics to big data, we will be able to get an insight into events that have yet to take place or behaviours that have yet to manifest.

It stands to reason that the more data collected, the more accurate the analyses and the more evidence-based decision made, thus it is vital to ensure that our capacity to collect data doesn't outrun our capability to analyse what we have collected.

It is almost impossible to think of a business area or industry that could not benefit from proper and appropriate analysis of big data and a number of industries have already been able to use the concept of big data to their benefit for example, scientific research including the human genome project and the Large Hadron Collider

(LHC) experiments. Multinational private sector companies are investing in improving their capacity to generate, handle and meaningfully analyse big data, with the public sector very keen to follow suit.

What does big data/predictive analytics mean to the O.R. community?

I believe that the increasing focus on big data is a huge potential opportunity for O.R. practitioners to have a considerable impact on analysis that informs the decision-making process, rather than a threat to the profession. The breadth of the O.R. toolkit, and the synergies that exist between the area and predictive analytics, puts us as O.R. students and practitioners at the forefront of the big data debate, as well as making us one of the most likely groups to produce meaningful analyses from such large and dynamic datasets. Combining skills such as problem structuring, modelling and stakeholder engagement with analytics can add meaningful value in terms of utilising big data.

The notion that data scientists will be the people to best get to grips with analytics and big data has become increasing popular, with the definition of a data scientist's skillset described as follows:
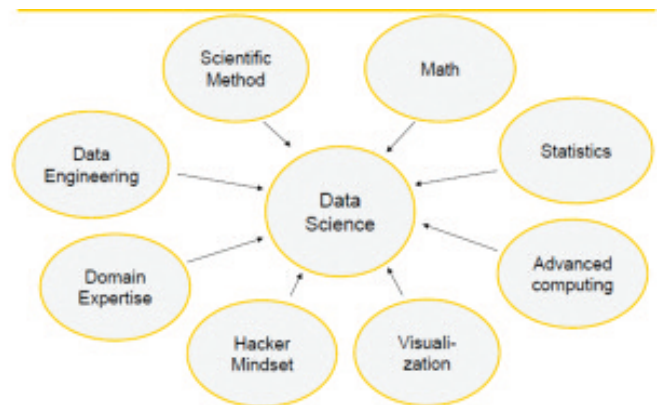
Taken from an Ernst & Young presentation 'How O.R. and Analytics Skills can be mutually beneficial.

It is easy to see from the skills mentioned above that O.R. analysts use the majority, if not all of these skills in their day-to-day work.

How does big data affect those studying O.R. or early O.R. careers, if at all?

Big data will undoubtedly play a part in the role of an operational researcher in the not so distant future; therefore it is important to always be aware of the landscape in which our roles currently exist. It is vital that we understand what big data is and how the

techniques in our toolkit can be modified to best accommodate it. It is also important that we are willing to be creative and flexible in our thinking, to ensure that we are solution driven rather than technique driven. There are some techniques that will be obviously applicable to big data and predictive analytics, as well as some which would not normally fall with the O.R. remit.

According to IBM, the global IT market will require 20 per cent more skilled big data analytics roles than it currently has within the next five years, hence they have spent a total of $20bn in the past year on acquisitions to allow them to 'move into a world of decision-making excellence with big data'. The number of possible avenues where big data could apply are almost infinite. A characteristic of operational researchers is that we have the ability to learn and use new techniques very quickly and so I think it would be wise for those of us just starting out in O.R. to ensure that we try to develop at least basic skills in predictive analytics and working with big data– it is also very likely to further enhance our employability prospects in the future.

It would be great to hear your thoughts on how you think big data and predictive analytics will change the O.R. landscape, and if you are already combining your O.R. skills with work in the area then drop me an email as I would be very interested in your viewpoint.

My email address is Louise.Maynard-Atem@dh.gsi.gov.uk. I look forward to hearing from you.

<OR>

YOUR 2014 TRAINING GUIDE

With this issue of Inside O.R.

Your 2014 Training Guide detailing 31 OR Society approved training courses.

The OR Society is committed to supporting you and ensuring that you have the tools you need to develop your career in line with changing industry needs.

You can also go online at

http://www.theorsociety.com/Pages/Training/TrainingCourses.aspx

to download your Training Guide or if you'd like more hard copies contact Jennie Phelps on

+44(0)121 234 7818 or by email at jennie.phelps@theorsociety.com

<OR>